

*Submitted to the Annals of Probability*

## LIMITS OF ONE DIMENSIONAL DIFFUSIONS

BY GEORGE LOWTHER

In this paper we look at the properties of limits of a sequence of real valued inhomogeneous diffusions. When convergence is only in the sense of finite-dimensional distributions then the limit does not have to be a diffusion. However, we show that as long as the drift terms satisfy a Lipschitz condition and the limit is continuous in probability, then it will lie in a class of processes that we refer to as the *almost-continuous diffusions*. These processes are strong Markov and satisfy an ‘almost-continuity’ condition. We also give a simple condition for the limit to be a continuous diffusion.

These results contrast with the multidimensional case where, as we show with an example, a sequence of two dimensional martingale diffusions can converge to a process that is both discontinuous and non-Markov.

**1. Introduction.** Suppose that we have a sequence of one dimensional diffusions, and that their finite-dimensional distributions converge. The aim of this paper is to show that, under a Lipschitz condition for the drift components of the diffusions, then the limit will lie in a class of processes that is an extension of the class of diffusions, which we refer to as the *almost-continuous diffusions*. Furthermore, we give a simple condition on this limit in order for it to be a continuous diffusion.

One way that an inhomogeneous diffusion can be defined is by an SDE

$$(1) \quad dX_t = \sigma(t, X_t) dW_t + b(t, X_t) dt$$

where  $W$  is a Brownian motion. Under certain conditions on  $\sigma$  and  $b$ , such as Lipschitz continuity, then it is well known that this SDE will have a unique solution (see [5] V.3, [6] IX.2, [7] V.11). Furthermore, whenever the solution is unique then  $X$  will be a strong Markov process (see [5] V.6, [7] V.21). More generally, we can consider all possible real valued and continuous strong Markov processes.

We now ask the question, if we have a sequence  $X^n$  of such processes whose finite-dimensional distributions converge, then does the limit have to be a continuous and strong Markov process? In general, the answer is no. There is no reason that the limit should either be continuous or be strong

---

*AMS 2000 subject classifications:* Primary 60J60, 60G44; secondary 60F99

*Keywords and phrases:* diffusion, martingale, strong Markov, finite-dimensional distributions

Markov. In the case of tight sequences (under the topology of locally uniform convergence) then convergence of the finite-dimensional distributions is enough to guarantee convergence under the weak topology, and the limits of continuous processes under the weak topology are themselves continuous (see [6] Chapter XIII or [1] Chapter 15).

However, we shall look at the case where the finite-dimensional distributions converge, but do not place any tightness conditions on the processes. In fact, we shall only place a Lipschitz condition on the increasing part of  $b(t, x)$  (w.r.t.  $x$ ) for processes given by the SDE (1), and place no conditions at all on  $\sigma(t, x)$ . We further generalize to processes that do not necessarily satisfy an SDE such as (1), but only have to satisfy the strong Markov property and a continuity condition.

In this case there is no need for the limit of continuous processes to be continuous, as we shall see later in a simple example. However, in the main result of this paper, we show that as long as the limit is continuous in probability, then it will be strong Markov and satisfy a pathwise continuity condition — which we shall refer to as being *almost-continuous*. Furthermore, under simple conditions on the limit, then it can be shown to be a continuous process.

The extension of continuous one dimensional diffusions that we require is given by the *almost-continuous diffusions* that we originally defined in [3].

**DEFINITION 1.1.** *Let  $X$  be a real valued stochastic process. Then,*

1.  *$X$  is strong Markov if for every bounded, measurable  $g : \mathbb{R} \rightarrow \mathbb{R}$  and every  $t \in \mathbb{R}_+$  there exists a measurable  $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$  such that*

$$f(\tau, X_\tau) = \mathbb{E}[g(X_{\tau+t}) \mid \mathcal{F}_\tau]$$

*for every finite stopping time  $\tau$ .*

2.  *$X$  is almost-continuous if it is cadlag, continuous in probability and given any two independent, identically distributed cadlag processes  $Y, Z$  with the same distribution as  $X$  and for every  $s < t \in \mathbb{R}_+$  we have*

$$\mathbb{P}(Y_s < Z_s, Y_t > Z_t \text{ and } Y_u \neq Z_u \text{ for every } u \in (s, t)) = 0$$

3.  *$X$  is an almost-continuous diffusion if it is strong Markov and almost-continuous.*

We shall often abbreviate ‘almost-continuous diffusion’ to ACD. Note that the almost-continuous property simply means that  $Y - Z$  cannot change sign without passing through zero, which is clearly a property of continuous processes. In [3] we applied coupling methods to prove that conditional

expectations of functions of such processes satisfy particularly nice properties, such as conserving monotonicity and, in the martingale case, Lipschitz continuity and convexity. These methods were originally used by [2] in the case of diffusions that are a unique solution to the SDE (1). As the results in this paper show, almost-continuous diffusions arise naturally as limits of continuous diffusions, and our method of proof will also employ similar coupling methods. Furthermore, in a future paper we shall show that, subject to a Lipschitz constraint on the drift component, any almost-continuous diffusion is a limit of continuous diffusions (under the topology of convergence of finite-dimensional distributions).

We now recall that the weak topology on the probability measures on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  is the smallest topology making the map  $\mu \mapsto \mu(f)$  continuous for every bounded and continuous  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ . In particular, a sequence  $(\mu_n)_{n \in \mathbb{N}}$  of probability measures on  $\mathbb{R}^n$  converges weakly to a measure  $\mu$  if and only if

$$\mu_n(f) \rightarrow \mu(f)$$

for every bounded and continuous  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ .

Now, suppose that we have real valued stochastic processes  $(X^n)_{n \in \mathbb{N}}$  and  $X$ , possibly defined on different probability spaces. Then, for any subset  $S$  of  $\mathbb{R}_+$ , we shall say that  $X^n$  converges to  $X$  in the sense of finite-dimensional distributions on  $S$  if and only if for every finite subset  $\{t_1, t_2, \dots, t_d\}$  of  $S$  then the distributions of  $(X_{t_1}^n, X_{t_2}^n, \dots, X_{t_d}^n)$  converges weakly to the distribution of  $(X_{t_1}, X_{t_2}, \dots, X_{t_d})$ .

We shall use the space of cadlag real valued processes (Skorokhod space) on which to represent the probability measures, and use  $X$  to represent the coordinate process.

$$\begin{aligned} \mathbf{D} &= \{\text{cadlag functions } \omega : \mathbb{R}_+ \rightarrow \mathbb{R}\}, \\ X &: \mathbb{R}_+ \times \mathbf{D} \rightarrow \mathbb{R}, (t, \omega) \mapsto X_t(\omega) \equiv \omega(t), \\ \mathcal{F} &= \sigma(X_t : t \in \mathbb{R}_+), \\ \mathcal{F}_t &= \sigma(X_s : s \in [0, t]). \end{aligned}$$

Then,  $(\mathbf{D}, \mathcal{F})$  is a measurable space and  $X$  is a cadlag process adapted to the filtration  $\mathcal{F}_t$ .

With these definitions, a sequence  $\mathbb{P}_n$  of probability measures on  $(\mathbf{D}, \mathcal{F})$  converges to  $\mathbb{P}$  in the sense of finite-dimensional distributions on a set  $S \subseteq \mathbb{R}_+$  if and only if

$$\mathbb{E}_{\mathbb{P}_n}[f(X_{t_1}, X_{t_2}, \dots, X_{t_d})] \rightarrow \mathbb{E}_{\mathbb{P}}[f(X_{t_1}, X_{t_2}, \dots, X_{t_d})]$$

as  $n \rightarrow \infty$  for every finite  $\{t_1, t_2, \dots, t_d\} \subseteq S$  and every continuous and bounded  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ .

We now state the main result of this paper.

**THEOREM 1.2.** *Let  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  be a sequence of probability measures on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  is an almost-continuous diffusion. Suppose that there exists a  $K \in \mathbb{R}$  such that, for every  $n \in \mathbb{N}$ , the process  $X$  decomposes as*

$$X_t = M_t^n + \int_0^t b_n(s, X_s) ds$$

where  $M^n$  is an  $\mathcal{F}$ -local martingale under  $\mathbb{P}_n$  and  $b_n : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$  is locally integrable and satisfies

$$b_n(t, y) - b_n(t, x) \leq K(y - x)$$

for every  $x < y \in \mathbb{R}$  and every  $t \in \mathbb{R}_+$ .

If  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions on a dense subset of  $\mathbb{R}_+$ , and  $X$  is continuous in probability under  $\mathbb{P}$ , then it is an almost-continuous diffusion under  $\mathbb{P}$ .

The proof of this result will be left until Sections 3 and 4. Note that in the special case where  $K = 0$  then the condition simply says that  $b_n(t, x)$  is decreasing in  $x$ . Furthermore, Theorem 1.2 reduces to the following simple statement in the martingale case.

**COROLLARY 1.3.** *Let  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  be a sequence of probability measures on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  is an ACD martingale. If  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions on a dense subset of  $\mathbb{R}_+$  and  $X$  is continuous in probability under  $\mathbb{P}$ , then it is an almost-continuous diffusion under  $\mathbb{P}$ .*

We can also give a simple condition on the measure  $\mathbb{P}$  from Theorem 1.2 and Corollary 1.3 in order for  $X$  to be a continuous process. Recall that the support of the real valued random variable  $X_t$  is the smallest closed subset  $C$  of the real numbers such that  $\mathbb{P}(X_t \in C) = 1$ .

**LEMMA 1.4.** *Let  $X$  be an almost-continuous process. If the support of  $X_t$  is connected for every  $t$  in  $\mathbb{R}_+$  outside of a countable set then  $X$  is continuous.*

The proof of this result is left until the end of Section 3, and follows quite easily from the properties of the *marginal support* of a process, which we studied in [3].

The results above (Theorem 1.2 and Corollary 1.3) are, in a sense, best possible. Certainly, it is possible for a continuous and Markov (but non-strong Markov) martingale to converge to a process that is neither almost-continuous nor Markov. Similarly, a strong Markov but discontinuous martingale can converge to a process that is not Markov. Furthermore, these results do not extend in any obvious way to multidimensional diffusions — in Section 2 we shall construct an example of a sequence of continuous martingale diffusions taking values in  $\mathbb{R}^2$ , and which converge to a discontinuous and non-Markov process.

Now, suppose that we have any sequence of probability measures  $\mathbb{P}_n$  on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  is an almost-continuous diffusion. In order to apply Theorem 1.2 we would need to be able to pass to a subsequence whose finite-dimensional distributions converge. It is well known that if the sequence is tight with respect to the Skorokhod topology, then it is possible to pass to a subsequence that converges weakly with respect to this topology (see [6] Chapter XIII or [1] Chapter 15). We do not want to restrict ourselves to this situation. Fortunately, it turns out that under fairly weak conditions on  $X$  then it is possible to pass to a subsequence that converges in the sense of finite-dimensional distributions. This follows from the results in [4], where they consider convergence under a topology that is much weaker than the Skorokhod topology, but is still strong enough to give convergence of the finite-dimensional distributions in an almost-everywhere sense.

By ‘convergence almost everywhere’ in the statement of the result below, we mean that there is an  $S \subseteq \mathbb{R}_+$  such that  $\mathbb{R}_+ \setminus S$  has zero Lebesgue measure and  $\mathbb{P}_{n_k} \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions on  $S$ . In particular,  $S$  must be a dense subset of  $\mathbb{R}_+$ . Recall that we are working under the natural filtration  $\mathcal{F}$  on Skorokhod space  $(\mathbf{D}, \mathcal{F})$ .

**THEOREM 1.5.** *Let  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  be a sequence of probability measures on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  has the decomposition*

$$X = M^n + A^n,$$

*where  $M^n$  is a cadlag  $\mathbb{P}_n$ -martingale and  $A^n$  is an adapted process with locally finite variation. Suppose further that for every  $t \in \mathbb{R}_+$  the sequence*

$$\mathbb{E}_{\mathbb{P}_n}[|X_t|] + \mathbb{E}_{\mathbb{P}_n}\left[\int_0^t |dA_s^n|\right]$$

*is finite and bounded.*

*Then, there exists a subsequence  $(\mathbb{P}_{n_k})_{k \in \mathbb{N}}$  and a measure  $\mathbb{P}$  on  $(\mathbf{D}, \mathcal{F})$  such that  $\mathbb{P}_{n_k} \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions almost everywhere on  $\mathbb{R}_+$ .*

PROOF. We use the results from [4] for tightness under the pseudo-path topology of a sequence of processes with bounded conditional variation.

For every  $k \in \mathbb{N}$  define the process  $Y_t^k \equiv 1_{\{t < k\}} X_t$ . Then the conditional variation of  $Y^k$  under the measure  $\mathbb{P}_n$  satisfies

$$V_n(Y^k) \leq \mathbb{E}_{\mathbb{P}_n} [|X_k|] + \mathbb{E}_{\mathbb{P}_n} \left[ \int_0^k |dA_s^n| \right]$$

which is bounded over all  $n \in \mathbb{N}$  by some constant  $L_k$ . Now define the process

$$Z_t = \sum_{k=1}^{\infty} 2^{-k} (L_k + 1)^{-1} Y_t^k = \theta(t) X_t$$

where  $\theta$  is the cadlag function

$$\theta(t) = \sum_{k=1}^{\infty} 2^{-k} (L_k + 1)^{-1} 1_{\{t < k\}}.$$

Then, the conditional variation of  $Z$  satisfies

$$V_n(Z) \leq \sum_{k=1}^{\infty} 2^{-k} (L_k + 1)^{-1} V_n(Y^k) \leq \sum_{k=1}^{\infty} 2^{-k} = 1.$$

So, by Theorem 4 of [4], there exists a subsequence  $(\mathbb{P}_{n_k})_{k \in \mathbb{N}}$  under which the laws of the process  $Z$  converge weakly (w.r.t. the pseudo-path topology) to the law of  $Z$  under a probability measure  $\mathbb{P}$ . Then by Theorem 5 of [4], we can pass to a further subsequence such that the finite-dimensional distributions of  $Z$  converge almost everywhere to those under  $\mathbb{P}$ . Finally, as  $X_t = \theta(t)^{-1} Z_t$ , we see that the finite-dimensional distributions of  $X$  also converge almost everywhere.  $\square$

These results give us a general technique that can be used to construct almost-continuous diffusions whose finite-dimensional distributions satisfy a desired property. That is, we first construct a sequence of almost-continuous diffusions whose distributions satisfy the required property in the limit. Then, we can appeal to Theorem 1.5 in order to pass to a convergent subsequence and use Theorem 1.2 to show that the limit is an almost-continuous diffusion. This is a method that we shall use in a later paper in order to construct ACD martingales with prescribed marginal distributions.

**2. Examples.** We give examples demonstrating how the convergence described in Theorem 1.2 behaves, and in particular show how a continuous diffusion can converge to a discontinuous process satisfying the almost-continuity condition.

We then give an example showing that Theorem 1.2 and Corollary 1.3 do not extend to multidimensional diffusions.

**2.1. Convergence to a reflecting Brownian motion.** We construct a simple example of continuous martingale diffusions converging to a reflecting Brownian motion. Consider the SDE

$$(2) \quad dX_t^n = \sigma(X_t^n) dW_t, \\ \sigma_n(x) = \max(1, -nx)$$

for each  $n \in \mathbb{N}$ , with  $X_0^n = 0$ . Here,  $W$  is a standard Brownian motion. As  $\sigma_n$  are Lipschitz continuous functions, these SDEs have a unique solution and  $X^n$  will be strong Markov martingales. In particular, they will be almost-continuous diffusions. We shall show that they converge to a reflecting Brownian motion.

The SDE (2) can be solved by a time change method, where we first choose any Brownian motion  $B$  and define the processes

$$(3) \quad A_t^n = \int_0^t \sigma_n(B_s)^{-2} ds, \\ T_t^n = \inf \{T \in \mathbb{R}_+ : A_T^n > t\}.$$

Then the process

$$X_t^n = B_{T_t^n}$$

gives a weak solution to SDE (2). We can take limits as  $n \rightarrow \infty$ ,

$$A_t^n \rightarrow A_t \equiv \int_0^t 1_{\{B_s \geq 0\}} ds.$$

If we now use  $A$  to define the time change,

$$(4) \quad T_t = \inf \{T \in \mathbb{R}_+ : A_T > t\}, \\ X_t = B_{T_t}$$

then  $X$  is a Brownian motion with the negative excursions removed, and so is a reflecting Brownian motion.

For every  $t \in \mathbb{R}_+$  we have  $X_t > 0$  (a.s.) and so  $A$  is strictly increasing in a neighbourhood of  $t$ . Therefore,  $T_t^n \rightarrow T_t$ . This shows that  $X_t^n \rightarrow X_t$

(a.s.), so the processes  $X^n$  do indeed converge to  $X$  in the sense of finite-dimensional distributions. However, it does not converge weakly with respect to the topology of locally uniform convergence. In fact, the minimum of  $X^n$  over any interval does not converge weakly to the minimum of  $X$ .

$$\inf_{s \leq t} X_s^n = \inf_{s \leq T_t^n} B_s \rightarrow \inf_{s \leq T_t} B_s < 0 = \inf_{s \leq t} X_s$$

for every  $t > 0$ . This example shows that a limit of martingale diffusions need not be a martingale. However, note that the support of  $X_t$  is  $[0, \infty)$  for any positive time  $t$ , and  $X$  has no drift over any interval that it does not hit 0. This is true more generally — whenever a process is a limit of one dimensional martingale diffusions, then it will behave like a local martingale except when it hits the edge of its support.

**2.2. Convergence to a symmetric Poisson process.** We show how continuous diffusions can converge to a discontinuous process, such as the symmetric Poisson process. By ‘symmetric Poisson process’ with rate  $\lambda$ , we mean a process with independent increments whose jumps occur according to a standard Poisson process with rate  $\lambda$  and such that the jump sizes are independent and take the values 1 and  $-1$ , with positive and negative jumps equally likely. Alternatively, it is the difference of two independent Poisson processes with rate  $\lambda/2$ .

If  $X$  is a symmetric Poisson process with  $X_0 = 0$ , then it follows that the support of  $X_t$  is  $\mathbb{Z}$  for every positive time  $t$  and it is easy to show that it satisfies the almost-continuous property.

We now let  $\sigma_n : \mathbb{R} \rightarrow \mathbb{R}$  be positive Lipschitz continuous functions such that  $\sigma_n^{-2}$  converges to a sum of delta functions at each integer point. For example, set

$$(5) \quad \sigma_n(x) = (\pi/n)^{1/4} \left( \sum_{k=-\infty}^{\infty} \exp(-n(x+k)^2) \right)^{-1/2}$$

In particular, this gives

$$(6) \quad \int f(x) \sigma_n(x)^{-2} dx \rightarrow \sum_{k \in \mathbb{Z}} f(k)$$

as  $n \rightarrow \infty$ , for all continuous functions  $f$  with compact support. We now consider the SDE

$$(7) \quad dX_t^n = \sigma_n(X_t^n) dW_t$$

where  $W$  is a standard Brownian motion, and  $X_0^n = 0$ . As  $\sigma_n$  is Lipschitz continuous,  $X^n$  will be an ACD martingale. We can solve this SDE by using a time changed Brownian motion, in the same way as for the previous example. So, let  $B$  be a standard Brownian motion and  $A_t^n, T_t^n$  be defined by equations (3). Then  $X_t^n = B_{T_t^n}$  solves SDE (7).

If we let  $L_t^a$  be the semimartingale local time of  $B$  at  $a$ , then it is jointly continuous in  $t$  and  $a$  and Tanaka's formula gives

$$A_t^n = \int L_t^a \sigma_n(a)^{-2} da.$$

Equation (6) allows us to take the limit as  $n$  goes to infinity,

$$A_t^n \rightarrow A_t \equiv \sum_{a \in \mathbb{Z}} L_t^a.$$

Then,  $A$  will be constant over any time interval for which  $B \notin \mathbb{Z}$  and it follows that if we define  $T_t$  and the time changed process  $X$  by equations (4) then the support of  $X_t$  will be contained in  $\mathbb{Z}$  for every time  $t$ . In fact,  $X$  will be a symmetric Poisson process.

As in the previous example, we have  $T_t^n \rightarrow T_t$  as  $n \rightarrow \infty$  (a.s.). Therefore  $X_t^n \rightarrow X_t$  (a.s.) for every  $t \in \mathbb{R}_+$ , showing that the continuous martingale diffusions converge to the discontinuous process  $X$ .

**2.3. A discontinuous and non-Markov limit of multidimensional martingale diffusions.** We give an example of a sequence of 2-dimensional continuous diffusions converging to a discontinuous and non-Markov process. This shows that Theorem 1.2 and Corollary 1.3 do not extend to the multidimensional case in any obvious way. To construct our example, first define the Lipschitz continuous function  $f : \mathbb{R} \rightarrow \mathbb{R}$  by

$$f(x) = \min \{|x - k| : k \in \mathbb{Z}\}.$$

Now let  $U$  be a normally distributed random variable with mean 0 and variance 1 (any random variable with support equal to  $\mathbb{R}$  and absolutely continuous distribution will do). Also, let  $\sigma_n$  be as in the previous example, defined by equation (5). Consider the SDE

$$\begin{aligned} dY_t^n &= f(nZ_t^n)\sigma_n(Y_t^n) dW_t, \\ dZ_t^n &= 0 \end{aligned}$$

where  $Y_0^n = 0$  and  $Z_0^n = U$ , and  $W$  is a standard Brownian motion. As  $f(nx)\sigma_n(x)$  is Lipschitz continuous, the processes  $(Y^n, Z^n)$  will be strong Markov martingales.

It is easy to solve this SDE. Let  $X^n$  be the processes defined in the previous example. Then a solution is given by

$$\begin{aligned} Y_t^n &= f(nU)X_t^n, \\ Z_t^n &= U. \end{aligned}$$

From the previous example, we know that  $X^n$  converges to a symmetric Poisson process  $X$  in the sense of finite-dimensional distributions. Also, from the definition of  $f$ ,  $f(nU)$  will converge weakly to the uniform distribution on  $[0, 1]$ . So, let  $V$  be a random variable uniformly distributed on  $[0, 1]$  and suppose that  $X$ ,  $V$  and  $U$  are independent. Setting

$$\begin{aligned} Y_t &= V X_t, \\ Z_t &= U \end{aligned}$$

then  $(Y^n, Z^n) \rightarrow (Y, Z)$  in the sense of finite-dimensional distributions. This process is both discontinuous and non-Markov, showing that the results of this paper do not extend to two dimensional processes.

In fact, I conjecture that for  $d > 1$  any  $d$ -dimensional cadlag stochastic process is a limit of martingale diffusions in the sense of finite-dimensional distributions, and for  $d > 2$  any such process is a limit of homogeneous martingale diffusions.

**3. Almost-Continuity.** We split the proof of Theorem 1.2 into two main parts. First, in this section, we show that the limit is an almost-continuous process, and we leave the proof that it is strong Markov until later. The main result that we shall prove in this section is the following.

**LEMMA 3.1.** *Let  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  be a sequence of probability measures on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  is an almost-continuous diffusion. If  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions on a dense subset of  $\mathbb{R}_+$  and  $X$  is continuous in probability under  $\mathbb{P}$ , then it is almost-continuous under  $\mathbb{P}$ .*

The method we use will be to reformulate the pathwise ‘almost-continuity’ property into a condition on the finite distributions of  $X$ . The idea is that given real numbers (or more generally, subsets of the reals)  $x < y$  and  $x' < y'$  then a coupling argument can be used to show that the probability of  $X$  going from  $x$  to  $y'$  multiplied by the probability of going from  $y$  to  $x'$  across a time interval  $[s, t]$  is bounded by the probability of going from  $x$  to  $x'$  multiplied by the probability of going from  $y$  to  $y'$ . The precise statement is as follows.

LEMMA 3.2. *Let  $\mathbb{P}$  be a probability measure on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  is continuous in probability. Then each of the following statements implies the next.*

1.  *$X$  is an almost-continuous diffusion.*
2. *for every  $s < t \in \mathbb{R}_+$ , non-negative  $\mathcal{F}_s$ -measurable random variables  $U, V$ , and real numbers  $a$  and  $b < c \leq d < e$ , then*

$$(8) \quad \begin{aligned} & \mathbb{E} [U \mathbf{1}_{\{X_s < a, d < X_t < e\}}] \mathbb{E} [V \mathbf{1}_{\{X_s > a, b < X_t < c\}}] \\ & \leq \mathbb{E} [U \mathbf{1}_{\{X_s < a, b < X_t < c\}}] \mathbb{E} [V \mathbf{1}_{\{X_s > a, d < X_t < e\}}]. \end{aligned}$$

3.  *$X$  is almost-continuous.*

We shall split the proof of this lemma into several parts. The approach that we use is to consider two independent copies of  $X$  and look at the first time that they cross. So, we start by defining the probability space on which these processes exist, which is just the product of  $(\mathbf{D}, \mathcal{F})$  with itself.

$$(9) \quad \begin{aligned} \mathbf{D}^2 &= \mathbf{D} \times \mathbf{D}, \\ \mathcal{F}^2 &= \mathcal{F} \otimes \mathcal{F}. \end{aligned}$$

Then we let  $Y$  and  $Z$  be the coordinate processes,

$$(10) \quad \begin{aligned} Y, Z : \mathbb{R}_+ \times \mathbf{D}^2 &\rightarrow \mathbb{R}, \\ Y_t(\omega_1, \omega_2) &\equiv X_t(\omega_1) = \omega_1(t), \\ Z_t(\omega_1, \omega_2) &\equiv X_t(\omega_2) = \omega_2(t). \end{aligned}$$

We also write  $\mathcal{F}_t^2$  for the filtration generated by  $Y$  and  $Z$ , which is just the product of  $\mathcal{F}_t$  with itself.

$$(11) \quad \mathcal{F}_t^2 \equiv \mathcal{F}_t \otimes \mathcal{F}_t = \sigma(Y_s, Z_s : s \in [0, t]).$$

Given any probability measure  $\mathbb{P}$  on  $(\mathbf{D}, \mathcal{F})$  we denote the measure on  $(\mathbf{D}^2, \mathcal{F}^2)$  formed by the product of  $\mathbb{P}$  with itself by  $\tilde{\mathbb{P}}$ .

$$(12) \quad \tilde{\mathbb{P}} \equiv \mathbb{P} \otimes \mathbb{P}.$$

In what follows, the notation  $\tilde{\mathbb{E}}[\cdot]$  will be used to denote expectations with respect to the measure  $\tilde{\mathbb{P}}$ . From these definitions,  $Y$  and  $Z$  are adapted cadlag processes, and under  $\tilde{\mathbb{P}}$  they are independent and identically distributed each with the same distribution as  $X$  has under  $\mathbb{P}$ . We now rewrite statement 2 of Lemma 3.2 in terms of the finite distributions of  $Y$  and  $Z$ .

LEMMA 3.3. *Given any probability measure  $\tilde{\mathbb{P}}$  on  $(\mathbf{D}, \mathcal{F})$ , statement 2 of Lemma 3.2 is equivalent to the statement that for every  $s < t \in \mathbb{R}_+$  and real numbers  $b < c \leq d < e$  then,*

$$(13) \quad \begin{aligned} & \tilde{\mathbb{P}}(Y_s < Z_s, b < Z_t < c, d < Y_t < e | \mathcal{F}_s^2) \\ & \leq \tilde{\mathbb{P}}(Y_s < Z_s, b < Y_t < c, d < Z_t < e | \mathcal{F}_s^2). \end{aligned}$$

PROOF. First, suppose that inequality (13) holds. Choose  $s < t \in \mathbb{R}_+$  and real numbers  $a$  and  $b < c \leq d < e$ . Also choose non-negative  $\mathcal{F}_s$ -measurable random variables  $U = u(X)$  and  $V = v(X)$ . Then the definition (12) of  $\tilde{\mathbb{P}}$  together with inequality (13) gives

$$\begin{aligned} & \mathbb{E}[U 1_{\{X_s < a, d < X_t < e\}}] \mathbb{E}[V 1_{\{X_s > a, b < X_t < c\}}] \\ &= \tilde{\mathbb{E}}[u(Y)v(Z) 1_{\{Y_s < a < Z_s\}} 1_{\{d < Y_t < e\}} 1_{\{b < Z_t < c\}}] \\ &= \tilde{\mathbb{E}}[u(Y)v(Z) 1_{\{Y_s < a < Z_s\}} \tilde{\mathbb{P}}(Y_s < Z_s, b < Z_t < c, d < Y_t < e | \mathcal{F}_s^2)] \\ &\leq \tilde{\mathbb{E}}[u(Y)v(Z) 1_{\{Y_s < a < Z_s\}} \tilde{\mathbb{P}}(Y_s < Z_s, b < Y_t < c, d < Z_t < e | \mathcal{F}_s^2)] \\ &= \tilde{\mathbb{E}}[u(Y)v(Z) 1_{\{Y_s < a < Z_s\}} 1_{\{b < Y_t < c\}} 1_{\{d < Z_t < e\}}] \\ &= \mathbb{E}[U 1_{\{X_s < a, b < X_t < c\}}] \mathbb{E}[V 1_{\{X_s > a, d < X_t < e\}}] \end{aligned}$$

as required.

Conversely, suppose that statement 2 of Lemma 3.2 holds. Now choose  $s < t \in \mathbb{R}_+$ , real numbers  $a' < a$  and  $b < c \leq d < e$  and bounded non-negative  $\mathcal{F}_s$ -measurable random variables  $U = u(X)$  and  $V = v(X)$ . Defining the  $\mathcal{F}_s^2$ -measurable random variable  $W = u(Y)v(Z)$ , then the definition (12) of  $\tilde{\mathbb{P}}$  together with inequality (8) gives

$$\begin{aligned} & \tilde{\mathbb{E}}[W 1_{\{a' \leq Y_s < a < Z_s, b < Z_t < c, d < Y_t < e\}}] \\ &= \mathbb{E}\left[\left(u(X) 1_{\{a' \leq X_s\}}\right) 1_{\{X_s < a, d < X_t < e\}}\right] \mathbb{E}\left[v(X) 1_{\{a < X_s\}} 1_{\{b < X_t < c\}}\right] \\ &\leq \mathbb{E}\left[\left(u(X) 1_{\{a' \leq X_s\}}\right) 1_{\{X_s < a, b < X_t < c\}}\right] \mathbb{E}\left[v(X) 1_{\{a < X_s\}} 1_{\{d < X_t < e\}}\right] \\ &= \tilde{\mathbb{E}}[W 1_{\{a' \leq Y_s < a < Z_s, b < Y_t < c, d < Z_t < e\}}]. \end{aligned}$$

For any  $\epsilon > 0$  we can set  $a' = (n - 1)\epsilon$  and  $a = n\epsilon$  in this inequality and

sum over  $n$ ,

$$\begin{aligned}
& \tilde{\mathbb{E}} \left[ W \mathbf{1}_{\{\exists n \in \mathbb{Z} \text{ s.t. } Y_s < n\epsilon < Z_s, b < Z_t < c, d < Y_t < e\}} \right] \\
&= \sum_{n=-\infty}^{\infty} \tilde{\mathbb{E}} \left[ W \mathbf{1}_{\{(n-1)\epsilon \leq Y_s < n\epsilon < Z_s, b < Z_t < c, d < Y_t < e\}} \right] \\
&\leq \sum_{n=-\infty}^{\infty} \tilde{\mathbb{E}} \left[ W \mathbf{1}_{\{(n-1)\epsilon \leq Y_s < n\epsilon < Z_s, b < Y_t < c, d < Z_t < e\}} \right] \\
&= \tilde{\mathbb{E}} \left[ W \mathbf{1}_{\{\exists n \in \mathbb{Z} \text{ s.t. } Y_s < n\epsilon < Z_s, b < Y_t < c, d < Z_t < e\}} \right]
\end{aligned}$$

Letting  $\epsilon$  decrease to 0 and using bounded convergence gives

$$(14) \quad \tilde{\mathbb{E}} \left[ W \mathbf{1}_{\{Y_s < Z_s, b < Z_t < c, d < Y_t < e\}} \right] \leq \tilde{\mathbb{E}} \left[ W \mathbf{1}_{\{Y_s < Z_s, b < Y_t < c, d < Z_t < e\}} \right].$$

Finally, we note that the set of bounded and non-negative  $\mathcal{F}_s^2$ -measurable random variables  $W$  for which inequality (14) holds is closed under taking positive linear combinations, and under taking increasing and decreasing limits. Therefore, inequality (14) holds for all bounded and non-negative  $\mathcal{F}_s^2$ -measurable random variables  $W$ , and inequality (13) follows from this.  $\square$

Using this result, it is now easy to prove that the first statement of Lemma 3.2 implies the second. The idea is to look at the processes  $Y$  and  $Z$  up until the first time that they touch, which is similar to the coupling method used in [2] to investigate the conditional expectations of convex functions of a martingale diffusion.

**LEMMA 3.4.** *If  $\mathbb{P}$  is a probability measure on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  is an almost-continuous diffusion then statement 2 of Lemma 3.2 holds.*

**PROOF.** First choose real numbers  $b < c \leq d < e$ , times  $s < t \in \mathbb{R}_+$ , and set

$$g_1(x) = \mathbf{1}_{\{b < x < c\}}, \quad g_2(x) = \mathbf{1}_{\{d < x < e\}}.$$

Then, by the strong Markov property, there exist measurable functions  $f_1, f_2 : [0, t] \times \mathbb{R} \rightarrow \mathbb{R}$  such that

$$\mathbf{1}_{\{\tau \leq t\}} f_i(\tau, X_\tau) = \mathbf{1}_{\{\tau \leq t\}} \mathbb{E}[g_i(X_t) | \mathcal{F}_\tau]$$

for  $i = 1, 2$  and for every stopping time  $\tau$ . This follows easily from definition 1.1 of the strong Markov property (see [3], Lemma 2.1). Furthermore it then

follows that

$$\begin{aligned} 1_{\{\tau \leq t\}} f_i(\tau, Y_\tau) &= 1_{\{\tau \leq t\}} \tilde{\mathbb{E}} [g_i(Y_t) | \mathcal{F}_\tau^2] \\ 1_{\{\tau \leq t\}} f_i(\tau, Z_\tau) &= 1_{\{\tau \leq t\}} \tilde{\mathbb{E}} [g_i(Z_t) | \mathcal{F}_\tau^2] \end{aligned}$$

for every  $\mathcal{F}^2$ -stopping time  $\tau$  (see [3], Lemma 2.2).

Now let  $\tau$  be the following stopping time.

$$\tau = \begin{cases} \inf \{u \in [s, \infty) : Y_u \geq Z_u\}, & \text{if } Y_s < Z_s, \\ \infty, & \text{otherwise.} \end{cases}$$

Strictly speaking,  $\tau$  will only be a stopping time with respect to the universal completion of the filtration. So, throughout this section we assume that all  $\sigma$ -algebras are replaced by their universal completions. Note that if  $Y_s < Z_s$  and  $\tau > t$  then  $Y_t < Z_t$  so  $g_1(Z_t)g_2(Y_t) = 0$ . Therefore

$$\begin{aligned} \tilde{\mathbb{P}}(Y_s < Z_s, b < Z_t < c, d < Y_t < e | \mathcal{F}_s^2) &= \tilde{\mathbb{E}} [1_{\{\tau \leq t\}} g_1(Z_t)g_2(Y_t) | \mathcal{F}_s^2] \\ &= \tilde{\mathbb{E}} [1_{\{\tau \leq t\}} f_1(\tau, Z_\tau) f_2(\tau, Y_\tau) | \mathcal{F}_s^2]. \end{aligned}$$

However, by almost-continuity, we have  $Y_\tau = Z_\tau$  whenever  $\tau < \infty$  ( $\tilde{\mathbb{P}}$  a.s.). So,

$$\begin{aligned} &\tilde{\mathbb{P}}(Y_s < Z_s, b < Z_t < c, d < Y_t < e | \mathcal{F}_s^2) \\ &= \tilde{\mathbb{E}} [1_{\{\tau \leq t\}} f_1(\tau, Y_\tau) f_2(\tau, Z_\tau) | \mathcal{F}_s^2] \\ &= \tilde{\mathbb{E}} [1_{\{\tau \leq t\}} g_1(Y_t)g_2(Z_t) | \mathcal{F}_s^2] \\ &= \tilde{\mathbb{P}}(\tau \leq t, b < Y_t < c, d < Z_t < e | \mathcal{F}_s^2) \\ &\leq \tilde{\mathbb{P}}(Y_s < Z_s, b < Y_t < c, d < Z_t < e | \mathcal{F}_s^2) \end{aligned}$$

The result now follows from Lemma 3.3.  $\square$

To prove that the second statement of Lemma 3.2 implies the third, we shall look at what happens when the processes  $Y$  and  $Z$  first cross after any given time. The idea is to show that they cannot jump past each other at this time, and therefore will be equal. As this will be a stopping time we start by rewriting statement 2 of Lemma 3.2 in terms of the distribution at a stopping time.

LEMMA 3.5. Let  $\tilde{\mathbb{P}}$  be a probability measure on  $(\mathbf{D}, \mathcal{F})$  such that statement 2 of Lemma 3.2 holds.

Let  $b < c \leq d < e$  be real numbers and set  $V = (b, c) \times (d, e)$ . Also let  $U$  be an open subset of  $\{(x, y) \in \mathbb{R}^2 : x < y\}$  that is disjoint from  $V$ , and for any  $\mathcal{F}^2$ -stopping time  $S$  define the stopping time

$$(15) \quad \tau_S^U = \begin{cases} \inf \{t > S : (Y_t, Z_t) \notin U\}, & \text{if } S < \infty \text{ and } (Y_S, Z_S) \in U, \\ \infty, & \text{otherwise.} \end{cases}$$

Then,

$$\tilde{\mathbb{P}}(\tau_S^U < \infty, (Z_{\tau_S^U}, Y_{\tau_S^U}) \in V) \leq \tilde{\mathbb{P}}(\tau_S^U < \infty, (Y_{\tau_S^U}, Z_{\tau_S^U}) \in V).$$

PROOF. Let  $t_{n,k} = k/n$  for all  $k \in \mathbb{Z}_{\geq 0}$  and  $n \in \mathbb{N}$ , and set

$$A_{n,k} = \left\{ S \leq t_{n,k} < \tau_S^U, (Y_S, Z_S) \in U \right\} \in \mathcal{F}_{t_{n,k}}^2.$$

We now let  $T_n$  be the stopping time

$$T_n = \inf \left\{ t_{n,k} : k \in \mathbb{N}, t_{n,k} \geq \tau_S^U > t_{n,k-1} \geq S \right\}$$

so that  $T_n \downarrow \tau_S^U$  as  $n \rightarrow \infty$ . Then we can apply Lemma 3.3,

$$(16) \quad \begin{aligned} \tilde{\mathbb{P}}(T_n < \infty, (Z_{T_n}, Y_{T_n}) \in V) &= \sum_{k=1}^{\infty} \tilde{\mathbb{P}}(T_n = t_{n,k}, (Z_{t_{n,k}}, Y_{t_{n,k}}) \in V) \\ &= \sum_{k=1}^{\infty} \tilde{\mathbb{P}}(A_{n,k-1} \cap \{(Z_{t_{n,k}}, Y_{t_{n,k}}) \in V\}) \\ &\leq \sum_{k=1}^{\infty} \tilde{\mathbb{P}}(A_{n,k-1} \cap \{(Y_{t_{n,k}}, Z_{t_{n,k}}) \in V\}) \\ &= \sum_{k=1}^{\infty} \tilde{\mathbb{P}}(T_n = t_{n,k}, (Y_{t_{n,k}}, Z_{t_{n,k}}) \in V) \\ &= \tilde{\mathbb{P}}(T_n < \infty, (Y_{T_n}, Z_{T_n}) \in V) \end{aligned}$$

Now suppose that  $\tau_S^U < \infty$  and  $(Z_{\tau_S^U}, Y_{\tau_S^U}) \in V$ . As  $T_n \downarrow \tau_S^U$  as  $n \rightarrow \infty$ , the right-continuity of  $Y$  and  $Z$  gives  $(Z_{T_n}, Y_{T_n}) \in V$  for large  $n$ . So, by bounded convergence

$$(17) \quad \tilde{\mathbb{P}}(\tau_S^U < \infty, (Z_{\tau_S^U}, Y_{\tau_S^U}) \in V) \leq \liminf_{n \rightarrow \infty} \tilde{\mathbb{P}}(T_n < \infty, (Z_{T_n}, Y_{T_n}) \in V).$$

Similarly, suppose that  $(Y_{T_n}, Z_{T_n}) \in V$  for infinitely many  $n$ . By the right-continuity of  $Y$  and  $Z$ , this gives  $b \leq Y_{\tau_S^U} \leq c$  and  $d \leq Z_{\tau_S^U} \leq e$ . So,

$$\limsup_{n \rightarrow \infty} 1_{\{T_n < \infty, (Y_{T_n}, Z_{T_n}) \in V\}} \leq 1_{\{\tau_S^U < \infty, b \leq Y_{\tau_S^U} \leq c, d \leq Z_{\tau_S^U} \leq e\}}.$$

Then monotone convergence gives

$$(18) \quad \begin{aligned} & \limsup_{n \rightarrow \infty} \tilde{\mathbb{P}}(T_n < \infty, (Y_{T_n}, Z_{T_n}) \in V) \\ & \leq \tilde{\mathbb{P}}(\tau_S^U < \infty, b \leq Y_{\tau_S^U} \leq c, d \leq Z_{\tau_S^U} \leq e). \end{aligned}$$

Combining inequalities (16), (17) and (18) gives

$$(19) \quad \begin{aligned} & \tilde{\mathbb{P}}(\tau_S^U < \infty, (Z_{\tau_S^U}, Y_{\tau_S^U}) \in V) \\ & \leq \tilde{\mathbb{P}}(\tau_S^U < \infty, b \leq Y_{\tau_S^U} \leq c, d \leq Z_{\tau_S^U} \leq e). \end{aligned}$$

Finally, set  $b_n = b + 1/n$ ,  $c_n = c - 1/n$ ,  $d_n = d + 1/n$  and  $e_n = e - 1/n$  for every  $n \in \mathbb{N}$ . Then inequality (19) with  $(b_n, c_n) \times (d_n, e_n)$  in place of  $V$  gives

$$\begin{aligned} & \tilde{\mathbb{P}}(\tau_S^U < \infty, (Z_{\tau_S^U}, Y_{\tau_S^U}) \in V) \\ &= \lim_{n \rightarrow \infty} \tilde{\mathbb{P}}(\tau_S^U < \infty, b_n < Z_{\tau_S^U} < c_n, d_n < Y_{\tau_S^U} < e_n) \\ &\leq \limsup_{n \rightarrow \infty} \tilde{\mathbb{P}}(\tau_S^U < \infty, b_n \leq Y_{\tau_S^U} \leq c_n, d_n \leq Z_{\tau_S^U} \leq e_n) \\ &= \tilde{\mathbb{P}}(\tau_S^U < \infty, (Y_{\tau_S^U}, Z_{\tau_S^U}) \in V). \square \end{aligned}$$

We shall use Lemma 3.5 to prove almost-continuity by showing that the probability of  $Y$  jumping from strictly below  $Z$  to above it is bounded by the probability of them jumping simultaneously. The following simple result will tell us that  $Y$  and  $Z$  cannot jump simultaneously.

**LEMMA 3.6.** *Let  $Y$  and  $Z$  be independent cadlag processes such that  $Y$  is continuous in probability. Then, with probability 1,  $Y_{t-} = Y_t$  or  $Z_{t-} = Z_t$  for every  $t > 0$ .*

**PROOF.** As  $Y$  is cadlag, there exist  $Y$ -measurable random times  $(S_n)_{n \in \mathbb{N}}$  such that  $\cup_{n \in \mathbb{N}} [S_n]$  contains all the jump times of  $Y$  almost-surely (see [1] Theorem 3.32). Without loss of generality, we may suppose that  $Y_{S_n-} \neq Y_{S_n}$  whenever  $S_n < \infty$ . Similarly, there exist  $Z$ -measurable random times

$(T_n)_{n \in \mathbb{N}}$  such that  $\cup_{n \in \mathbb{N}} [T_n]$  contains all the jump times of  $Z$  almost-surely, and such that  $Z_{T_n-} \neq Z_{T_n}$  whenever  $T_n < \infty$ . Then,

$$\mathbb{P}(\exists t \in \mathbb{R}_+ \text{ s.t. } Y_{t-} \neq Y_t \text{ and } Z_{t-} \neq Z_t) \leq \sum_{m,n=1}^{\infty} \mathbb{P}(S_m = T_n < \infty).$$

However, the independence of  $S_m$  and  $T_n$  together with the continuity in probability of  $Y$  gives

$$\begin{aligned} \mathbb{P}(S_m = T_n < \infty) &= \sum_{t \in \mathbb{R}_+} \mathbb{P}(S_m = t) \mathbb{P}(T_n = t) \\ &\leq \sum_{t \in \mathbb{R}_+} \mathbb{P}(Y_{t-} \neq Y_t) \mathbb{P}(Z_{t-} \neq Z_t) \\ &= 0. \square \end{aligned}$$

This simple result together with Lemma 3.5 gets us some way towards showing that  $X$  is almost-continuous.

LEMMA 3.7. *Let  $\mathbb{P}$  be a probability measure on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  is continuous in probability, and such that statement 2 of Lemma 3.2 holds. Then*

$$\tilde{\mathbb{P}}(\exists t > 0 \text{ s.t. } Y_{t-} < Z_t < Y_t) = 0.$$

PROOF. Choose any real numbers  $b < c < d$  and let  $U, V$  be the sets

$$\begin{aligned} U &= (-\infty, b) \times (b, c), \\ V &= (b, c) \times (c, d). \end{aligned}$$

Then  $U \cap V = \emptyset$ , so letting  $\tau_s^U$  be the stopping time given by equation (15) for any  $s \in \mathbb{R}_+$ , we can apply Lemma 3.5 to get

$$(20) \quad \tilde{\mathbb{P}}\left(\tau_s^U < \infty, (Z_{\tau_s^U}, Y_{\tau_s^U}) \in V\right) \leq \tilde{\mathbb{P}}\left(\tau_s^U < \infty, (Y_{\tau_s^U}, Z_{\tau_s^U}) \in V\right).$$

However, if  $\tau_s^U < \infty$  and  $(Y_{\tau_s^U}, Z_{\tau_s^U}) \in V$  then  $Y_{\tau_s^U} > b \geq Y_{\tau_s^U-}$  and  $Z_{\tau_s^U} > c \geq Z_{\tau_s^U-}$ . By Lemma 3.6 the processes  $Y$  and  $Z$  cannot jump simultaneously, so this has zero probability. Inequality (20) then gives

$$(21) \quad \tilde{\mathbb{P}}\left(\tau_s^U < \infty, (Z_{\tau_s^U}, Y_{\tau_s^U}) \in V\right) = 0.$$

Now suppose that  $(Y_{t-}, Z_{t-}) \in U$  and  $(Z_t, Y_t) \in V$  for some time  $t$ . Then, by left-continuity, there exists an  $s < t$  such that  $s \in \mathbb{Q}_+$  and  $(Y_s, Z_s) \in U$

for every  $u \in [s, t]$ . In this case  $\tau_s^U = t$ . Therefore equation (21) gives

$$\begin{aligned}
& \tilde{\mathbb{P}}(\exists t \in \mathbb{R}_+ \text{ s.t. } Y_{t-} < b < Z_{t-} = Z_t < c < Y_t < d) \\
& \leq \tilde{\mathbb{P}}(\exists t \in \mathbb{R}_+ \text{ s.t. } (Y_{t-}, Z_{t-}) \in U, (Z_t, Y_t) \in V) \\
(22) \quad & \leq \sum_{s \in \mathbb{Q}_+} \tilde{\mathbb{P}}(\tau_s^U < \infty, (Z_{\tau_s^U}, Y_{\tau_s^U}) \in V) \\
& = 0.
\end{aligned}$$

Note that for every  $t$  such that  $Y_{t-} < Z_t < Y_t$  then Lemma 3.6 tells us that  $Z_{t-} = Z_t$ . So, by (22)

$$\begin{aligned}
& \tilde{\mathbb{P}}(\exists t \in \mathbb{R}_+ \text{ s.t. } Y_{t-} < Z_t < Y_t) \\
& = \tilde{\mathbb{P}}(\exists t \in \mathbb{R}_+ \text{ s.t. } Y_{t-} < Z_{t-} = Z_t < Y_t) \\
& \leq \sum_{a < b < c < d \in \mathbb{Q}} \tilde{\mathbb{P}}(\exists t \in \mathbb{R}_+ \text{ s.t. } Y_{t-} < b < Z_{t-} = Z_t < c < Y_t < d) \\
& = 0. \square
\end{aligned}$$

Lemma 3.7 shows that  $Y$  cannot jump from strictly below to strictly above  $Z$ . However, it does not rule out the possibility that  $Y$  can approach  $Z$  from below, then jump to above  $Z$  (which would contradict almost-continuity). In order to show that this behaviour is not possible, we shall again make use of Lemma 3.5. The idea is to reduce it to showing that it is not possible for  $Y$  to approach  $Z$  from below, then jump downwards. In fact this behaviour is ruled out by the conclusion of Lemma 3.7, but it is far from obvious that this is the case. We shall make use of some results that we proved in [3]. First, we restate the definition of the marginal support used in [3].

**DEFINITION 3.8.** *Let  $X$  be a real valued stochastic process. Then, its marginal support is*

$$\text{MSupp}(X) = \{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : x \in \text{Supp}(X_t)\}.$$

As we showed in [3], the marginal support of a process  $X$  is Borel measurable, and the relevance of the marginal support to our current argument is given by the following result.

**LEMMA 3.9.** *If  $X$  is a cadlag real valued process which is continuous in probability then the following are equivalent.*

1. *The set*

$$\{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : X_{t-} < x < X_t\}.$$

*is disjoint from  $\text{MSupp}(X)$  with probability one.*

2. Given two independent cadlag processes  $Y$  and  $Z$ , each with the same distribution as  $X$ , then

$$\mathbb{P}(\exists t > 0 \text{ s.t. } Y_{t-} < Z_t < Y_t) = 0$$

PROOF. See [3], Lemma 4.7.  $\square$

We also make use of the following result, which says that it is not possible for  $Y$  to approach  $Z$  from below and then jump downwards to a value strictly less than  $Z$ .

LEMMA 3.10. Let  $X$  be a cadlag real valued process which is continuous in probability, and such that the set

$$\{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : X_{t-} < x < X_t \text{ or } X_t < x < X_{t-}\}$$

is disjoint from  $\text{MSupp}(X)$  with probability one.

Also, let  $Y$  and  $Z$  be independent cadlag processes each with the same distribution as  $X$ . For any  $s \in \mathbb{R}_+$  let  $T$  be the random time

$$T = \begin{cases} \inf\{t \in \mathbb{R}_+ : t \geq s, Y_t \geq Z_t\}, & \text{if } Y_s < Z_s, \\ \infty, & \text{otherwise,} \end{cases}$$

and  $(T_n)_{n \in \mathbb{N}}$  be the random times

$$T_n = \begin{cases} \inf\{t \in \mathbb{R}_+ : t \geq s, Y_t + 1/n \geq Z_t\}, & \text{if } Y_s < Z_s, \\ \infty, & \text{otherwise.} \end{cases}$$

Then  $T_n \uparrow T$  as  $n \rightarrow \infty$  (a.s.). Also,  $T_n < T$  whenever  $T < \infty$  and  $Y_T \neq Z_T$  (a.s.).

PROOF. See [3], Lemma 4.9.  $\square$

We can now combine these results to prove Lemma 3.2. As we mentioned previously, the idea is to show that it is not possible for  $Y$  to approach  $Z$  from below and then jump past it.

PROOF. First, statement 1 implies statement 2 by Lemma 3.4. So, we now suppose that statement 2 holds. Then by Lemma 3.7 we have

$$\tilde{\mathbb{P}}(\exists t > 0 \text{ s.t. } Y_{t-} < Z_t < Y_t) = 0.$$

Applying Lemma 3.9 shows that the set

$$\{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : X_{t-} < x < X_t\}$$

is disjoint from  $\text{MSupp}(X)$  with probability one. Similarly, we can apply the same argument to  $-X$  to see that

$$\{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : X_t < x < X_{t-}\}$$

is also disjoint from  $\text{MSupp}(X)$  with probability one. Therefore, the requirements of Lemma 3.10 are satisfied. For any  $s \in \mathbb{R}_+$  and  $n \in \mathbb{N}$  let  $T$  and  $T_n$  be the stopping times defined by Lemma 3.10. Also, define the stopping times

$$S_n = \begin{cases} T_n, & \text{if } T_n < T, \\ \infty, & \text{otherwise,} \end{cases}$$

$$S = \begin{cases} T, & \text{if } T_n < T \text{ for every } n \in \mathbb{N}, \\ \infty, & \text{otherwise.} \end{cases}$$

By Lemma 3.10,  $T_n \uparrow T$ , and so  $S_n \uparrow\uparrow S$  whenever  $S < \infty$ . Now let  $A$  be the set

$$A = \left\{ a \in \mathbb{R} : \tilde{\mathbb{P}}(S < \infty, Z_{S-} = a) = 0 \right\}.$$

As  $\mathbb{R} \setminus A$  is countable, we see that  $A$  is a dense subset of  $\mathbb{R}$ . We now choose any  $b < c < d \in A$  and set

$$U = \{(x, y) \in \mathbb{R} : x < y < c\},$$

$$V = (b, c) \times (c, d).$$

Now fix any  $t > s$  and let  $S'_n$  be the stopping time,

$$S'_n = \begin{cases} S_n, & \text{if } S_n \leq t, \\ \infty, & \text{otherwise.} \end{cases}$$

Also let  $\tau_{S'_n}^U$  be the stopping time defined by equation (15). Then, by Lemma 3.5

$$(23) \quad \tilde{\mathbb{P}}\left(\tau_{S'_n}^U < \infty, (Z_{\tau_{S'_n}^U}, Y_{\tau_{S'_n}^U}) \in V\right) \leq \tilde{\mathbb{P}}\left(\tau_{S'_n}^U < \infty, (Y_{\tau_{S'_n}^U}, Z_{\tau_{S'_n}^U}) \in V\right).$$

Also, if  $S < \infty$  then  $S = T$  so, by the definition of  $T$  we have  $Y_S \geq Z_S$ . So,  $(Y_S, Z_S) \notin U$ . Now consider the following cases,

- $S \leq t$  and  $Z_{S-} < c$ . Then, as  $S_n \uparrow\uparrow S$ , we see that  $\tau_{S'_n}^U = S$  for large  $n$ .
- $S \leq t$  and  $Z_{S-} > c$ . Then  $Z_{S_n} > c$  for large  $n$  and so  $\tau_{S'_n}^U = \infty$ .
- $S > t$ . Then  $S'_n = \infty$  for large  $n$ .

The case where  $Z_{S-} = c$  is ruled out because we chose  $c \in A$ . Therefore, we can take the limit as  $n$  goes to infinity in inequality (23),

$$\begin{aligned} & \tilde{\mathbb{P}}(S \leq t, Z_{S-} < c, (Z_S, Y_S) \in V) \\ & \leq \tilde{\mathbb{P}}(S \leq t, Z_{S-} < c, (Y_S, Z_S) \in V) \end{aligned}$$

As  $Y_S \geq Z_S$ , the right hand side of this inequality is 0,

$$\tilde{\mathbb{P}}(S \leq t, Z_{S-} < c, b < Z_S < c < Y_S < d) = 0.$$

Therefore, letting  $B$  be any countable and dense subset of  $A$ ,

$$\begin{aligned} & \tilde{\mathbb{P}}(S \leq t, Z_{S-} = Z_S < Y_S) \\ & \leq \sum_{b < c < d \in B} \tilde{\mathbb{P}}(S \leq t, b < Z_{S-} = Z_S < c < Y_S < d) \\ & = 0. \end{aligned}$$

This shows that it is not possible for  $Y$  to approach  $Z$  from below, then jump upwards to above  $Z$ . Similarly, replacing  $(Y, Z)$  by  $(-Z, -Y)$  in the above argument gives

$$\tilde{\mathbb{P}}(S \leq t, Z_S < Y_S = Y_{S-}) = 0.$$

Lemma 3.6 says that  $Z_{S-} = Z_S$  or  $Y_{S-} = Y_S$  whenever  $S < \infty$ ,

$$\begin{aligned} \tilde{\mathbb{P}}(S \leq t, Z_S < Y_S) &= \tilde{\mathbb{P}}(S \leq t, Z_{S-} = Z_S < Y_S) \\ &\quad + \tilde{\mathbb{P}}(S \leq t, Z_S < Y_S = Y_{S-}) \\ &= 0. \end{aligned}$$

That is,  $Y_S = Z_S$  whenever  $S \leq t$  (a.s.). Finally, from the statement of Lemma 3.10, we know that  $Y_T = Z_T$  whenever  $T_n = T < \infty$ . So,  $Y_T = Z_T$  whenever  $T \leq t$ . So whenever  $Y_s < Z_s$  and  $Z_t < Y_t$  we have  $s < T < t$  and  $Z_T = Y_T$ . Therefore

$$\tilde{\mathbb{P}}(Y_s < Z_s, Y_t > Z_t \text{ and } Y_u \neq Z_u \text{ for every } u \in (s, t)) = 0. \quad \square$$

We now move on to the proof of Lemma 3.1. We start off by considering the case where the finite distributions converge everywhere, rather than just on a dense subset of  $\mathbb{R}_+$ .

**LEMMA 3.11.** *Let  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  be probability measures on  $(\mathbf{D}, \mathcal{F})$  which satisfy property 2 of Lemma 3.2. If  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions, then  $\mathbb{P}$  also satisfies this property.*

PROOF. First choose  $s < t \in \mathbb{R}_+$  and real numbers  $a$  and  $b < c \leq d < e$ . Also choose times  $t_1, t_2, \dots, t_r \in [0, s]$  and non-negative continuous and bounded functions  $u, v : \mathbb{R}^r \rightarrow \mathbb{R}$ . We let  $U, V$  be the  $\mathcal{F}_s$ -measurable random variables

$$\begin{aligned} U &= u(X_{t_1}, X_{t_2}, \dots, X_{t_r}), \\ V &= v(X_{t_1}, X_{t_2}, \dots, X_{t_r}). \end{aligned}$$

Then, for any real numbers  $a_1 < a_2$  and  $b' < c' \leq d' < e'$  inequality (8) gives

$$\begin{aligned} &\mathbb{E}_{\mathbb{P}_n} \left[ U \mathbf{1}_{\{X_s < a_1, d' < X_t < e'\}} \right] \mathbb{E}_{\mathbb{P}_n} \left[ V \mathbf{1}_{\{X_s > a_2, b' < X_t < c'\}} \right] \\ &= \mathbb{E}_{\mathbb{P}_n} \left[ \left( U \mathbf{1}_{\{X_s < a_1\}} \right) \mathbf{1}_{\{X_s < a_2, d' < X_t < e'\}} \right] \mathbb{E}_{\mathbb{P}_n} \left[ V \mathbf{1}_{\{X_s > a_2, b' < X_t < c'\}} \right] \\ &\leq \mathbb{E}_{\mathbb{P}_n} \left[ \left( U \mathbf{1}_{\{X_s < a_1\}} \right) \mathbf{1}_{\{X_s < a_2, b' < X_t < c'\}} \right] \mathbb{E}_{\mathbb{P}_n} \left[ V \mathbf{1}_{\{X_s > a_2, d' < X_t < e'\}} \right] \\ &= \mathbb{E}_{\mathbb{P}_n} \left[ U \mathbf{1}_{\{X_s < a_1, b' < X_t < c'\}} \right] \mathbb{E}_{\mathbb{P}_n} \left[ V \mathbf{1}_{\{X_s > a_2, d' < X_t < e'\}} \right] \end{aligned}$$

If we take limits as  $n$  goes to infinity and use convergence of the finite-dimensional distributions then this gives

$$\begin{aligned} &\mathbb{E}_{\mathbb{P}} \left[ U \mathbf{1}_{\{X_s < a_1, d' < X_t < e'\}} \right] \mathbb{E}_{\mathbb{P}} \left[ V \mathbf{1}_{\{X_s > a_2, b' < X_t < c'\}} \right] \\ &\leq \mathbb{E}_{\mathbb{P}} \left[ U \mathbf{1}_{\{X_s \leq a_1, b' \leq X_t \leq c'\}} \right] \mathbb{E}_{\mathbb{P}} \left[ V \mathbf{1}_{\{X_s \geq a_2, d' \leq X_t \leq e'\}} \right]. \end{aligned}$$

Taking limits as  $a_1 \uparrow a$ ,  $a_2 \downarrow a$ ,  $b' \downarrow b$ ,  $c' \uparrow c$ ,  $d' \downarrow d$  and  $e' \uparrow e$  gives

$$\begin{aligned} &\mathbb{E}_{\mathbb{P}} \left[ U \mathbf{1}_{\{X_s < a, d < X_t < e\}} \right] \mathbb{E}_{\mathbb{P}} \left[ V \mathbf{1}_{\{X_s > a, b < X_t < c\}} \right] \\ &\leq \mathbb{E}_{\mathbb{P}} \left[ U \mathbf{1}_{\{X_s < a, b < X_t < c\}} \right] \mathbb{E}_{\mathbb{P}} \left[ V \mathbf{1}_{\{X_s > a, d < X_t < e\}} \right]. \end{aligned}$$

Note that the set of pairs of random variables  $(U, V)$  for which this inequality is true is closed under bounded limits, and under increasing limits. Therefore, it extends to all non-negative and  $\mathcal{F}_s$ -measurable random variables  $(U, V)$ .  $\square$

We now extend this result to the case where convergence is on a dense subset of  $\mathbb{R}_+$ .

COROLLARY 3.12. *Let  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  be probability measures on  $(\mathbf{D}, \mathcal{F})$  which satisfy property 2 of Lemma 3.2. If  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions on a dense subset of  $\mathbb{R}_+$ , then  $\mathbb{P}$  also satisfies this property.*

PROOF. Let  $S$  be a dense subset of  $\mathbb{R}_+$  such that  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions on  $S$ . Then, for every  $m \in \mathbb{N}$  we can find a sequence  $(t_{m,k})_{k \in \mathbb{N}} \in S$  such that  $(k-1)/m \leq t_{m,k} < k/m$ . We define

$$\begin{aligned}\theta_m : \mathbb{R}_+ &\rightarrow S, \\ \theta_m(t) &= \min \{t_{m,k} : k \in \mathbb{N}, t_{m,k} > t\}.\end{aligned}$$

Then,  $\theta_m$  is a right-continuous and non-decreasing function, so if we define the process  $X_t^m \equiv X_{\theta_m(t)}$  then it is clear that property 2 of Lemma 3.2 is satisfied if we replace  $X$  by  $X^m$  under the measure  $\mathbb{P}_n$  (and use the natural filtration generated by  $X^m$ ). Therefore, letting  $\mathbb{Q}_{n,m}$  be the measure on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  has the same distribution as  $X^m$  has under  $\mathbb{P}_n$ , then property 2 of Lemma 3.2 is satisfied for the measure  $\mathbb{Q}_{n,m}$ . Also, for every  $m \in \mathbb{N}$ , let  $\mathbb{Q}_m$  be the measure on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  has the same distribution as  $X^m$  has under  $\mathbb{P}$ .

As  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions on  $S$ , then it follows that

$$(24) \quad \mathbb{Q}_{n,m} \rightarrow \mathbb{Q}_m$$

as  $n \rightarrow \infty$ , in the sense of finite-dimensional distributions (on all of  $\mathbb{R}_+$ ). Also,  $\theta_m(t) \geq t$  and  $\theta_m(t) \rightarrow t$  as  $m \rightarrow \infty$ . Therefore, right-continuity of  $X_t$  gives  $X_t^m \rightarrow X_t$ . So,

$$(25) \quad \mathbb{Q}_m \rightarrow \mathbb{P}$$

in the sense of finite-dimensional distributions as  $m \rightarrow \infty$ . Applying Lemma 3.11 to the limit (24) tells us that  $\mathbb{Q}_m$  satisfies property 2 of Lemma 3.2. Finally, applying Lemma 3.2 to limit (25) shows that  $\mathbb{P}$  also satisfies this property.  $\square$

Now, we can finish off the proof of Lemma 3.1.

PROOF. As  $X$  is an almost-continuous diffusion under each of the measures  $\mathbb{P}_n$ , the second property of Lemma 3.2 is satisfied. Corollary 3.12 then tells us that  $\mathbb{P}$  also satisfies this property. So, Lemma 3.2 says that  $X$  is almost-continuous under  $\mathbb{P}$ .  $\square$

We shall now give a quick proof of Lemma 1.4. First, we will make use of the following result that says that the paths of a process lie inside its marginal support.

LEMMA 3.13. *Let  $X$  be a cadlag real valued process. Then, with probability 1, we have*

$$\{(t, X_t) : t \in \mathbb{R}_+\} \subseteq \text{MSupp}(X).$$

*Furthermore, if  $X$  is continuous in probability then*

$$\{(t, X_{t-}) : t \in \mathbb{R}_+\} \subseteq \text{MSupp}(X)$$

*with probability 1.*

PROOF. See [3], Lemma 4.3 and Lemma 4.4. □

The proof of Lemma 1.4 follows easily.

PROOF. First, by the statement of the lemma, there exists a countable  $S \subseteq \mathbb{R}_+$  such that  $\text{Supp}(X_t)$  is connected for every  $t \in \mathbb{R}_+ \setminus S$ . As  $X$  is cadlag, there exist stopping times  $(\tau_n)_{n \in \mathbb{N}}$  such that the jump times of  $X$  are almost-surely contained in  $\cup_{n \in \mathbb{N}} [\![\tau_n]\!]$  (see [1] Theorem 3.32). By Lemma 3.13 we have

$$(\tau_n, X_{\tau_n}), (\tau_n, X_{\tau_n-}) \in \text{MSupp}(X)$$

whenever  $\tau_n < \infty$  (a.s.). Also, by almost-continuity, the second condition of Lemma 3.9 is satisfied, and therefore the set

$$\{(\tau_n, x) : x \in \mathbb{R}, X_{\tau_n-} < x < X_{\tau_n}\}$$

is almost-surely disjoint from the marginal support of  $X$ , whenever  $\tau_n < \infty$ . So, the connected open components of the complement of the set

$$A = \{x \in \mathbb{R} : (\tau_n, x) \in \text{MSupp}(X)\}$$

includes the interval  $(X_{\tau_n-}, X_{\tau_n})$  whenever  $\tau_n < \infty$  and  $X_{\tau_n-} < X_{\tau_n}$ , so  $A$  is not connected in this case and we see that  $\tau_n \in S$ . Therefore, the continuity in probability of  $X$  gives

$$\mathbb{P}(\tau_n < \infty, X_{\tau_n-} < X_{\tau_n}) \leq \sum_{t \in S} \mathbb{P}(\tau_n = t, X_{t-} < X_t) = 0.$$

Similarly, applying the same argument to  $-X$ ,

$$\mathbb{P}(\tau_n < \infty, X_{\tau_n-} > X_{\tau_n}) = 0.$$

So  $X_{\tau_n-} = X_{\tau_n}$  (a.s.) whenever  $\tau_n < \infty$ , which shows that  $X$  is continuous. □

**4. The Strong Markov Property.** The aim of this section is to complete the proof of Theorem 1.2 by showing that the process  $X$  is strong Markov under the measure  $\mathbb{P}$ . To do this, we make use of the property that conditional expectations of Lipschitz continuous functions of  $X_t$  are themselves Lipschitz continuous. In this definition we write  $f'$  and  $g'$  to denote the derivatives  $df(x)/dx$  and  $dg(x)/dx$  in the measure-theoretic sense, which always exist for Lipschitz continuous functions.

**DEFINITION 4.1.** *Let  $X$  be any real valued and adapted stochastic process. We shall say that it satisfies the Lipschitz property if for all  $s < t \in \mathbb{R}_+$  and every bounded Lipschitz continuous  $g : \mathbb{R} \rightarrow \mathbb{R}$  with  $|g'| \leq 1$ , there exists a Lipschitz continuous  $f : \mathbb{R} \rightarrow \mathbb{R}$  with  $|f'| \leq 1$  and,*

$$f(X_s) = \mathbb{E}[g(X_t)|\mathcal{F}_s].$$

The reason that we use this property is that it is preserved under taking limits in the sense of finite-dimensional distributions (as we shall see), and is a sufficient condition for the process  $X$  to be strong Markov.

**LEMMA 4.2.** *Let  $X$  be a cadlag adapted real valued process that satisfies the Lipschitz property. Then, it is strong Markov.*

**PROOF.** Choose any  $s > 0$  and Lipschitz continuous and bounded  $g : \mathbb{R} \rightarrow \mathbb{R}$ . By the Lipschitz property there exists an  $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$  such that  $f(t, x)$  is Lipschitz continuous in  $x$  and

$$(26) \quad f(\tau, X_\tau) = \mathbb{E}[g(X_{\tau+s})|\mathcal{F}_\tau] \quad (\text{a.s.})$$

for every  $\tau \in \mathbb{R}_+$ . By linearity, this extends to all stopping times  $\tau$  that take only finitely many values in  $\mathbb{R}_+$ . We shall show that  $f(t, x)$  is right-continuous in  $t$  on the marginal support of  $X$ . So, pick any  $t \geq 0$  and sequence  $t_n \downarrow t$ . By the right-continuity of  $X$  and uniform continuity of  $f(t, x)$  and  $g(x)$  in  $x$ ,

$$\mathbb{E}[g(X_{t+s})|\mathcal{F}_{t+}] = \lim_{n \rightarrow \infty} f(t_n, X_{t_n}) = \lim_{n \rightarrow \infty} f(t_n, X_t)$$

where convergence is in probability. Taking conditional expectations with respect to  $\mathcal{F}_t$ ,

$$\lim_{n \rightarrow \infty} f(t_n, X_t) = \mathbb{E}[g(X_{t+s})|\mathcal{F}_t] = f(t, X_t).$$

By uniform continuity in  $x$ , this shows that  $f(t_n, x) \rightarrow f(t, x)$  for every  $x$  in the support of  $X_t$  and it follows that  $f(t, x)$  is right-continuous in  $t$  on

the marginal support of  $X$ . So Lemma 3.13 shows that  $f(t, X_t)$  is a right-continuous process and, by taking right limits in  $\tau$ , (26) extends to all finite stopping times  $\tau$ . Finally, the monotone class lemma extends this to all measurable and bounded  $g$ .  $\square$

We now prove the Lipschitz property for the case where the drift term  $b(t, x)$  is decreasing in  $x$ .

**LEMMA 4.3.** *Let  $\mathbb{P}$  be a probability measure on  $(\mathbf{D}, \mathcal{F})$  under which  $X$  is an almost-continuous diffusion which decomposes as*

$$X_t = M_t + \int_0^t b(s, X_s) ds$$

*where  $M$  is a local martingale and  $b : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$  is locally integrable such that  $b(t, x)$  is decreasing in  $x$ .*

*Then,  $X$  satisfies the Lipschitz property under  $\mathbb{P}$ .*

**PROOF.** Fix any  $s < t \in \mathbb{R}_+$  and let  $g : \mathbb{R} \rightarrow \mathbb{R}$  be bounded and Lipschitz continuous with  $|g'| \leq 1$ . As  $X$  is strong Markov, there exists a measurable  $h : [0, t] \times \mathbb{R} \rightarrow \mathbb{R}$  such that

$$1_{\{\tau \leq t\}} h(\tau, X_\tau) = 1_{\{\tau \leq t\}} \mathbb{E}[g(X_t) | \mathcal{F}_\tau]$$

for every stopping time  $\tau$ . This follows easily from the strong Markov property (see [3], Lemma 2.1).

We now let  $(\mathbf{D}^2, \mathcal{F}^2, (\mathcal{F}_t^2)_{t \in \mathbb{R}_+}, \tilde{\mathbb{P}})$  be the filtered probability space defined by equations (9), (11) and (12). We also let  $Y, Z$  be the stochastic processes on  $(\mathbf{D}^2, \mathcal{F}^2)$  defined by equation (10). Then,  $Y, Z$  are independent adapted cadlag processes each with the same distribution under  $\tilde{\mathbb{P}}$  as  $X$  has under  $\mathbb{P}$ . So they have the decompositions

$$(27) \quad \begin{aligned} Y_u &= M_u^1 + \int_0^u b(v, Y_v) dv, \\ Z_u &= M_u^2 + \int_0^u b(v, Z_v) dv \end{aligned}$$

for local martingales  $M^1, M^2$ . Furthermore,  $Y, Z$  will also be strong Markov and satisfy

$$(28) \quad \begin{aligned} 1_{\{\tau \leq t\}} h(\tau, Y_\tau) &= 1_{\{\tau \leq t\}} \tilde{\mathbb{E}}[g(Y_t) | \mathcal{F}_\tau], \\ 1_{\{\tau \leq t\}} h(\tau, Z_\tau) &= 1_{\{\tau \leq t\}} \tilde{\mathbb{E}}[g(Z_t) | \mathcal{F}_\tau] \end{aligned}$$

for all  $\mathcal{F}_s^2$ -stopping times  $\tau$ . This follows quite easily from the definitions of  $Y, Z$  (see [3], Lemma 2.2).

Now let  $\tau$  be the stopping time

$$\tau = \inf \{u \in [s, \infty) : Y_u \geq Z_u\}.$$

We note that  $\{\tau > s\} = \{Y_s < Z_s\}$ . Then equation (28) gives

$$\begin{aligned} 1_{\{\tau > s\}} (h(s, Z_s) - h(s, Y_s)) &= 1_{\{\tau > s\}} \tilde{\mathbb{E}} \left[ g(Z_t) - g(Y_t) \mid \mathcal{F}_s^2 \right] \\ &= \tilde{\mathbb{E}} \left[ 1_{\{\tau \geq t\}} (g(Z_t) - g(Y_t)) \mid \mathcal{F}_s^2 \right] \\ &\quad + \tilde{\mathbb{E}} \left[ 1_{\{t > \tau > s\}} (h(\tau, Z_\tau) - h(\tau, Y_\tau)) \mid \mathcal{F}_s^2 \right]. \end{aligned}$$

The almost-continuity of  $X$  gives  $Y_\tau = Z_\tau$  whenever  $t > \tau > s$ , so

$$\begin{aligned} (29) \quad 1_{\{\tau > s\}} (h(s, Z_s) - h(s, Y_s)) &= \tilde{\mathbb{E}} \left[ 1_{\{\tau \geq t\}} (g(Z_t) - g(Y_t)) \mid \mathcal{F}_s^2 \right] \\ &= \tilde{\mathbb{E}} \left[ 1_{\{\tau > s\}} (g(Z_{t \wedge \tau}) - g(Y_{t \wedge \tau})) \mid \mathcal{F}_s^2 \right] \\ &\leq \tilde{\mathbb{E}} \left[ 1_{\{\tau > s\}} (Z_{t \wedge \tau} - Y_{t \wedge \tau}) \mid \mathcal{F}_s^2 \right]. \end{aligned}$$

Here, we made use of the condition that  $|g'| \leq 1$ .

Let  $N$  be the local martingale

$$(30) \quad N_u = 1_{\{\tau > s\}} \left( M_{u \wedge \tau}^2 - M_{u \wedge \tau}^1 + \int_0^s (b(u, Z_u) - b(u, Y_u)) du \right)$$

defined over  $u \geq s$ . Then for every  $u \geq s$ , the condition that  $b(u, x)$  is decreasing in  $x$  gives

$$(31) \quad N_u - 1_{\{\tau > s\}} (Z_{u \wedge \tau} - Y_{u \wedge \tau}) = \int_s^{u \wedge \tau} (b(v, Y_v) - b(v, Z_v)) dv \geq 0.$$

So, substituting into inequality (29),

$$(32) \quad 1_{\{\tau > s\}} (h(s, Z_s) - h(s, Y_s)) \leq 1_{\{\tau > s\}} (Z_s - Y_s) + \mathbb{E} \left[ N_t - N_s \mid \mathcal{F}_s^2 \right].$$

Inequality (31) shows that  $N_u$  is a non-negative local martingale and is therefore a supermartingale. So  $\mathbb{E}[N_t | \mathcal{F}_s^2] \leq N_s$  and inequality (32) gives

$$h(s, Z_s) - h(s, Y_s) \leq Z_s - Y_s$$

whenever  $Y_s < Z_s$  (almost surely). Replacing  $h$  by  $-h$  in the above argument will also give the above inequality with  $Y$  and  $Z$  interchanged on the left

hand side. Furthermore, the inequality still holds if we interchange  $Y$  and  $Z$  on both sides (by symmetry). So

$$|h(s, Z_s) - h(s, Y_s)| \leq |Z_s - Y_s|$$

(almost surely). As  $Y_s, Z_s$  are independent and each have the same distribution as  $X_s$ , this shows that  $h(s, \cdot)$  is Lipschitz continuous in an almost-sure sense. That is, there is a measurable  $A \subseteq \mathbb{R}$  such that  $\mathbb{P}(X_s \in A) = 1$  and such that

$$|h(s, x) - h(s, y)| \leq |x - y|$$

for every  $x, y \in A$ . Therefore we can define  $f : \mathbb{R} \rightarrow \mathbb{R}$  by  $f(x) = h(s, x)$  for all  $x \in A$ . By uniform continuity, this extends uniquely to the closure  $\bar{A}$  of  $A$  such that

$$|f(x) - f(y)| \leq |x - y|$$

for every  $x, y \in \bar{A}$ . Then we can extend  $f$  linearly across each open interval in the complement of  $\bar{A}$  so that  $f$  is Lipschitz continuous with  $|f'| \leq 1$ . Finally,  $f(X_s) = h(s, X_s)$  whenever  $X_s \in A$  so,

$$f(X_s) = h(s, X_s) = \mathbb{E}[g(X_t) | \mathcal{F}_s]. \quad \square$$

We now extend this result to cover the case where  $b(t, x)$  just satisfies the Lipschitz condition required by Theorem 1.2.

**COROLLARY 4.4.** *Let  $X$  be an almost-continuous diffusion that decomposes as*

$$X_t = M_t + \int_0^t b(s, X_s) ds$$

*where  $M$  is a local martingale,  $b : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$  is locally integrable and such that there exists a  $K \in \mathbb{R}$  satisfying*

$$b(t, y) - b(t, x) \leq K(y - x)$$

*for every  $t \in \mathbb{R}_+$  and  $x < y \in \mathbb{R}$ .*

*Then,  $e^{-Kt} X_t$  satisfies the Lipschitz property.*

**PROOF.** If we set  $Y_t = e^{-Kt} X_t$  then  $Y$  is an almost-continuous diffusion and integration by parts gives

$$Y_t = N_t + \int_0^t c(s, Y_s) ds$$

where

$$N_t = X_0 + \int_0^t e^{-Ks} dM_s,$$

$$c(t, y) = e^{-Kt} b(t, e^{Kt} y) - Ky.$$

As  $N$  is a local martingale and  $c(t, y)$  is decreasing in  $y$ , the result follows from Lemma 4.3.  $\square$

In order to show that the limit in Theorem 1.2 is strong Markov, we shall show that  $e^{-Kt} X_t$  satisfies the Lipschitz property. This works because this property is preserved under taking limits in the sense of finite-dimensional distributions.

**LEMMA 4.5.** *Let  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  and  $\mathbb{P}$  be probability measures on  $(\mathbf{D}, \mathcal{F})$  such that  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions.*

*If the Lipschitz property for  $X$  is satisfied under each  $\mathbb{P}_n$  then it is also satisfied under  $\mathbb{P}$ .*

**PROOF.** Fix any  $s < t \in \mathbb{R}_+$  and any bounded and Lipschitz continuous  $g : \mathbb{R} \rightarrow \mathbb{R}$  such that  $|g| \leq K$  and  $|g'| \leq 1$ . Then, by the Lipschitz property for  $X$  under  $\mathbb{P}_n$ , there exist Lipschitz continuous functions  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  such that  $|f_n| \leq K$ ,  $|f'_n| \leq 1$  and

$$f_n(X_s) = \mathbb{E}_{\mathbb{P}_n}[g(X_t)|\mathcal{F}_s].$$

Now let  $S \subseteq \mathbb{R}$  be the support of  $X_s$  under  $\mathbb{P}$ . We shall show that  $f_n$  converges pointwise on  $S$  as  $n \rightarrow \infty$ . So, pick any  $x \in S$  and any  $\epsilon > 0$ . Let  $\theta : \mathbb{R} \rightarrow \mathbb{R}$  be any continuous and non-negative function with support contained in  $[x - \epsilon, x + \epsilon]$  such that  $\theta(x) > 0$ . As  $x \in S$  we have

$$\mathbb{E}_{\mathbb{P}}[\theta(X_s)] = \delta > 0.$$

We use the following simple identity

$$\begin{aligned} \delta (f_n(x) - f_m(x)) &= \mathbb{E}_{\mathbb{P}_n}[\theta(X_s)f_n(x)] - \mathbb{E}_{\mathbb{P}_m}[\theta(X_s)f_m(x)] \\ &\quad - (\mathbb{E}_{\mathbb{P}_n}[\theta(X_s)] - \mathbb{E}_{\mathbb{P}}[\theta(X_s)]) f_n(x) \\ &\quad + (\mathbb{E}_{\mathbb{P}_m}[\theta(X_s)] - \mathbb{E}_{\mathbb{P}}[\theta(X_s)]) f_m(x) \end{aligned}$$

Convergence of the distribution of  $X_s$  under  $\mathbb{P}_n$  to its distribution under  $\mathbb{P}$  (as  $n \rightarrow \infty$ ) tells us that the final two terms on the right hand side of this inequality vanish as we take limits, so

$$\begin{aligned} (33) \quad &\delta \limsup_{m,n \rightarrow \infty} |f_n(x) - f_m(x)| \\ &\leq \limsup_{m,n \rightarrow \infty} |\mathbb{E}_{\mathbb{P}_n}[\theta(X_s)f_n(x)] - \mathbb{E}_{\mathbb{P}_m}[\theta(X_s)f_m(x)]|. \end{aligned}$$

Also, Lipschitz continuity of  $f_n$  on the interval  $[x - \epsilon, x + \epsilon]$  gives

$$|\theta(X_s)(f_n(X_s) - f_n(x))| \leq \epsilon\theta(X_s)$$

so

$$\begin{aligned} & |\mathbb{E}_{\mathbb{P}_n}[\theta(X_s)f_n(x)] - \mathbb{E}_{\mathbb{P}_m}[\theta(X_s)f_m(x)]| \\ & \leq |\mathbb{E}_{\mathbb{P}_n}[\theta(X_s)f_n(X_s)] - \mathbb{E}_{\mathbb{P}_m}[\theta(X_s)f_m(X_s)]| \\ & \quad + \epsilon\mathbb{E}_{\mathbb{P}_n}[\theta(X_s)] + \epsilon\mathbb{E}_{\mathbb{P}_m}[\theta(X_s)] \\ & = |\mathbb{E}_{\mathbb{P}_n}[\theta(X_s)g(X_t)] - \mathbb{E}_{\mathbb{P}_m}[\theta(X_s)g(X_t)]| \\ & \quad + \epsilon\mathbb{E}_{\mathbb{P}_n}[\theta(X_s)] + \epsilon\mathbb{E}_{\mathbb{P}_m}[\theta(X_s)]. \end{aligned}$$

If we take limits as  $m, n \rightarrow \infty$  then the convergence of the finite-dimensional distributions of  $\mathbb{P}_n$  and  $\mathbb{P}_m$  to  $\mathbb{P}$  shows that the right hand side of this inequality converges to  $2\epsilon\delta$ ,

$$\limsup_{m,n \rightarrow \infty} |\mathbb{E}_{\mathbb{P}_n}[\theta(X_s)f_n(x)] - \mathbb{E}_{\mathbb{P}_m}[\theta(X_s)f_m(x)]| \leq 2\epsilon\delta.$$

Substituting this into inequality (33) gives

$$\limsup_{m,n \rightarrow \infty} |f_n(x) - f_m(x)| \leq 2\epsilon.$$

As this is true for every  $\epsilon > 0$ , the sequence  $f_n(x)$  is Cauchy, and therefore converges as  $n$  goes to infinity. So we can define  $f : S \rightarrow \mathbb{R}$  by  $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ . Then

$$(34) \quad |f(x) - f(y)| = \lim_{n \rightarrow \infty} |f_n(x) - f_n(y)| \leq |x - y|$$

for every  $x, y \in S$ . So  $f$  is Lipschitz continuous on  $S$ . By interpolating and extrapolating  $f$  linearly across the connected open components of  $\mathbb{R} \setminus S$ , we can extend it to a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  such that inequality (34) is satisfied for all  $x, y \in \mathbb{R}$ . So,  $f$  is Lipschitz continuous with  $|f'| \leq 1$ . Also, we can choose  $f$  such that  $|f| \leq K$ . To complete the proof of the lemma, it only remains to show that  $f(X_s) = \mathbb{E}_{\mathbb{P}}[g(X_t)|\mathcal{F}_s]$ .

Now set

$$h(x) = \limsup_{n \rightarrow \infty} |f_n(x) - f(x)|$$

so that  $h$  vanishes on  $S$ , and is a bounded Lipschitz continuous function satisfying  $|h| \leq 2K$  and  $|h'| \leq 2$ . By uniform continuity of the functions  $f_n$ , the convergence is uniform on bounded subsets of  $\mathbb{R}$ . That is, for every  $A > 0$ ,

$$|f_n(x) - f(x)| \leq h(x) + 1/A$$

for all large  $n$  and  $|x| \leq A$ . So,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}_n} [|f_n(X_s) - f(X_s)|] \\ & \leq \limsup_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}_n} [h(X_s)] + 1/A + 2K \limsup_{n \rightarrow \infty} \mathbb{P}_n (|X_s| > A) \\ & \leq \mathbb{E}_{\mathbb{P}} [h(X_s)] + 1/A + 2K \mathbb{P} (|X_s| \geq A) \\ & = 1/A + 2K \mathbb{P} (|X_s| \geq A). \end{aligned}$$

Letting  $A$  increase to infinity gives

$$(35) \quad \limsup_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}_n} [|f_n(X_s) - f(X_s)|] = 0.$$

Finally, choose any finite set of times  $t_1, t_2, \dots, t_d \in [0, s]$ , choose any bounded and continuous  $u : \mathbb{R}^d \rightarrow \mathbb{R}$ , and let  $U$  be the  $\mathcal{F}_s$ -measurable random variable

$$U = u(X_{t_1}, X_{t_2}, \dots, X_{t_d})$$

Then, letting  $L$  be an upper bound for  $|u|$ , we can use the equality  $f_n(X_s) = \mathbb{E}_{\mathbb{P}_n} [g(X_t) | \mathcal{F}_s]$  and equation (35) to get

$$\begin{aligned} |\mathbb{E}_{\mathbb{P}} [U(f(X_s) - g(X_t))]| &= \lim_{n \rightarrow \infty} |\mathbb{E}_{\mathbb{P}_n} [U(f(X_s) - g(X_t))]| \\ &\leq \limsup_{n \rightarrow \infty} |\mathbb{E}_{\mathbb{P}_n} [U(f_n(X_s) - g(X_t))]| \\ &\quad + L \limsup_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}_n} [|f_n(X_s) - f(X_s)|] \\ &= 0. \end{aligned}$$

Therefore,  $\mathbb{E}_{\mathbb{P}} [Uf(X_s)] = \mathbb{E}_{\mathbb{P}} [Ug(X_t)]$ . By the monotone class lemma this extends to all bounded and  $\mathcal{F}_s$ -measurable  $U$ , so  $f(X_s) = \mathbb{E}_{\mathbb{P}} [g(X_t) | \mathcal{F}_s]$ .  $\square$

This result is for convergence everywhere of the finite-dimensional distributions. It is easy to extend it to only require convergence on a dense subset of  $\mathbb{R}_+$ .

**COROLLARY 4.6.** *Let  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  and  $\mathbb{P}$  be probability measures on  $(\mathbf{D}, \mathcal{F})$  such that  $\mathbb{P}_n \rightarrow \mathbb{P}$  in the sense of finite-dimensional distributions on a dense subset of  $\mathbb{R}_+$ .*

*If the Lipschitz property for  $X$  is satisfied under each  $\mathbb{P}_n$  then it is also satisfied under  $\mathbb{P}$ .*

**PROOF.** We imply this result from Lemma 4.5 in the same way that Corollary 3.12 followed from Lemma 3.11. So, let  $\mathbb{Q}_{n,m}$  and  $\mathbb{Q}_m$  be the probability measures on  $(\mathbf{D}, \mathcal{F})$  defined in the proof of Corollary 3.12.

It is clear that the Lipschitz property for  $X$  under  $\mathbb{P}_n$  implies that it also satisfies the Lipschitz property under  $\mathbb{Q}_{n,m}$ . Then Lemma 4.5 applied to the limit (24) says that  $X$  satisfies the Lipschitz property under  $\mathbb{Q}_m$ . Applying Lemma 4.5 to the limit (25) gives the result.  $\square$

We finally prove Theorem 1.2.

PROOF. First, Lemma 3.1 says that  $X$  is almost-continuous under the measure  $\mathbb{P}$ . Also, Corollary 4.4 says that  $e^{-Kt}X_t$  satisfies the Lipschitz property under  $\mathbb{P}_n$ , so by Corollary 4.6 it also satisfies the Lipschitz property under  $\mathbb{P}$ . Lemma 4.2 then says that  $e^{-Kt}X_t$  is a strong Markov process under  $\mathbb{P}$ , and therefore  $X$  is also a strong Markov process.  $\square$

## REFERENCES

- [1] HE, S., WANG, J., AND YAN, J. (1992). *Semimartingale theory and stochastic calculus*. Kexue Chubanshe (Science Press), Beijing. [MRMR1219534](#)
- [2] HOBSON, D. G. (1998). Volatility misspecification, option pricing and superreplication via coupling. *Ann. Appl. Probab.* **8**, 1, 193–205. [MRMR1620358](#)
- [3] LOWTHER, G. (2008). Properties of expectations of functions of martingale diffusions. Pre-print available as arXiv:0801.0330v1 [math.PR] at arxiv.org.
- [4] MEYER, P.-A. AND ZHENG, W. A. (1984). Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré Probab. Statist.* **20**, 4, 353–372. [MRMR771895](#)
- [5] PROTTER, P. E. (2004). *Stochastic integration and differential equations*, Second ed. Applications of Mathematics (New York), Vol. **21**. Springer-Verlag, Berlin. Stochastic Modelling and Applied Probability. [MRMR2020294](#)
- [6] REVUZ, D. AND YOR, M. (1991). *Continuous martingales and Brownian motion*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. **293**. Springer-Verlag, Berlin. [MRMR1083357](#)
- [7] ROGERS, L. C. G. AND WILLIAMS, D. (1987). *Diffusions, Markov processes, and martingales*. Vol. 2. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York. [MRMR921238](#)